<u>AMENDMENTS TO THE CLAIMS:</u>

Please amend the claims as follows:

1. (Currently amended) A <u>computer-implemented</u> process for <u>generating</u>creating a filter for selecting documents, comprising:

identifying a plurality of profiles <u>from a set of training documents;</u>

creating a sub-filter for each of said plurality of profiles, each of said sub-filters having an input and an output<u>,</u>

connecting each of said inputs at a single node; and

combining each of said outputs <u>to thereby form said filter for selecting documents; and</u>

<u>storing said filter for selecting documents in a computer readable medium, said filter for selecting documents being accessible by computer readable program code for filtering and selecting documents;</u>

<u>wherein creating each sub-filter comprises the steps of</u>

<u>selecting a first subset and a second subset from a set of training documents,</u>

<u>extracting a set of features from the first subset, and</u>

<u>calculating a score threshold for the set of features from the second subset.</u>

2. (Currently amended) A <u>computer-implemented</u> process, as in claim 1, further comprising:

mutiplexing<u>multiplexing</u> said outputs of said sub-filters to create a first filter output.

3. (Currently amended) A <u>computer-implemented</u> process, as in claim 2, further comprising:

combining said first filter output with at least one sub-filter output to create at least one second filter output.

4. (Currently amended) A <u>computer-implemented</u> process, as in claim 3, further comprising:

aggregating said first filter output and said second filter output using a function.

Amendment

5. (Currently amended) A <u>computer-implemented</u> process, as in claim 4, wherein said function is a linear function.

6. (Currently amended) A <u>computer-implemented</u> process, as in claim 4, wherein said function is a weighted function.

7. (Currently amended) A <u>computer-implemented</u> process, as in claim 4, wherein said function is a ~~thresholded~~<u>threshold</u> aggregation function.

8. (Currently amended) A <u>computer-implemented</u> process for <u>generating</u>~~creating~~ an <u>ensemble</u> filter for selecting documents from a <u>document source</u>~~stream of documents,~~ comprising:

identifying a plurality of sets of documents from ~~said stream of documents~~ <u>a set of training documents</u>;

identifying a plurality of profiles corresponding to said plurality of sets of documents;

creating a plurality of sub-filters ~~using~~ <u>for</u> said plurality of profiles; ~~and~~

combining a first one of said plurality of sub-filters with a second one of said plurality of sub-filters to <u>thereby</u> create an ensemble filter<u>; and</u>

<u>storing said ensemble filter for selecting documents in a computer readable medium, said filter for selecting documents being accessible by computer readable program code for filtering and selecting documents;</u>

<u>wherein creating each sub-filter comprises the steps of</u>

<u>selecting a first subset and a second subset from a set of documents,</u>

<u>extracting a set of features from the first subset, and</u>

<u>calculating a score threshold for the set of features from the second subset.</u>

9 – 15 (Canceled)

16. (New) A computer-implemented process, as in claim 8, wherein the document source is a database of documents.

Amendment

17.   (New) A computer-implemented process, as in claim 8, wherein the document source is a stream of documents.

18.   (New) A computer-implemented process for generating a filtering profile for selecting documents comprising the steps of:

selecting a first subset of training documents and a second subset of training documents from a set of training documents;

creating a filtering profile having an input and an output by extracting a set of features from the first subset of training documents;

creating a score threshold for the filtering profile by calculating an optimal score threshold for the set of features from the second subset of training documents; and

storing said filtering profile for selecting documents in a computer readable medium, said filtering profile for selecting documents being accessible by computer readable program code for filtering and selecting documents.

19.   (New) The computer-implemented process of claim 18 wherein extracting a set of features further comprises identifying features in the first subset using a term frequency inverse document frequency ("TF-IDF") score.

20.   (New) The computer-implemented process of claim 18 wherein extracting a set of features further comprises performing a Rocchio algorithm.

21.   (New) The computer-implemented process of claim 18 wherein calculating an optimal score threshold further comprises performing a beta-gamma thresholding algorithm.

22.   (New) An apparatus for generating a document filter, comprising:
a processing system; and
a memory coupled to the processing system, wherein the processor is configured to:

select a first subset of training documents and a second subset of training documents from a set of training documents;

Amendment

create a filtering profile having an input and an output by extracting a set of features from the first subset of training documents;

create a score threshold for the filtering profile by calculating an optimal score threshold for the set of features from the second subset of training documents; and

store said filtering profile for selecting documents in a computer readable medium, said filtering profile for selecting documents being accessible by computer readable program code for filtering and selecting documents.

23.   (New) An article of manufacture comprising a computer readable medium having executable program code embodied therein for generating a document filter, wherein the executable program code is adapted to cause the processing system to:

select a first subset of training documents and a second subset of training documents from a set of training documents;

create a filtering profile having an input and an output by extracting a set of features from the first subset of training documents;

create a score threshold for the filtering profile by calculating an optimal score threshold for the set of features from the second subset of training documents; and

store said filtering profile for selecting documents in a computer readable medium, said filtering profile for selecting documents being accessible by computer readable program code for filtering and selecting documents.

Amendment